Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

V.

ON A THEOREM OF SYLVESTER'S RELATING TO NON-DEGENERATE* MATRICES.

By HENRY TABER, CLARK UNIVERSITY.

Presented March 9, 1892.

IN his *Memoir on Matrices,* Cayley enunciated the theorem that the most general matrix commutative with a given matrix was a rational integral function of it. Of this theorem Sylvester gave a proof, in the Johns Hopkins University Circulars, Vol. III. pp. 34 and 57, for the case in which the latent roots of the given matrix were all distinct, but pointed out that, when the latent roots of the given matrix were not all distinct, the theorem did not always hold. Subsequently, in the Comptes Rendus, Vol. XCVIII. p. 471, Sylvester stated that he had proved that Cayley's theorem held for a matrix of the third order which was not degenerate *(dérogatoire),*† irrespective of equalities between its latent roots; and he further stated that Cayley's theorem was probably true for non-degenerate matrices of any order; i. e. that the most general matrix commutative with a non-degenerate matrix of any order whatever was a rational integral function of it. I shall in this paper verify this conjecture of Sylvester's. The theorem follows very simply from one given by the author in a note in these Proceedings for May 26, 1891,‡ which contains the complete solution of the problem to find the most general matrix commutative with a given matrix.

The theorem is then to be proved for the case in which the latent roots of the given matrix are not all distinct. If the distinct latent roots of the given matrix Ω are g_1, g_2, g_3, etc., of multiplicity

---

* Sylvester employs the term *matrice dérogatoire* (Comptes Rendus, Vol. XCVIII. p. 471) to denote a matrix whose identical equation is of order lower than the order of the matrix; such a matrix I term a *degenerate* matrix.

† See preceding note.

‡ These Proceedings, Vol. XXVI. p 64.

m, n, p, etc., respectively, then Ω may be written as $\omega \theta \omega^{-1}$, where ω is a non-vacuous matrix, and

$$\theta = \left(\begin{array}{cccc} \boxed{\theta_1} & & & \\ & \boxed{\theta_2} & & \\ & & \boxed{\theta_3} & \\ & & & \ddots \end{array} \right),$$

in which all the constituents are zero except those in the squares θ_1, θ_2, θ_3, etc., along the principal diagonal; these correspond respectively to the latent roots g_1, g_2, g_3, etc., and are subsidiary matrices each of order equal to the multiplicity of the respective latent root to which it corresponds. Since Ω is supposed non-degenerate, the characteristics of its latent roots are $(m;\ 1, 1, \ldots .\ 1)$, $(n;\ 1, 1, \ldots .\ 1)$, $(p;\ 1, 1, \ldots .\ 1)$, etc., respectively. Moreover, since the characteristics of g_1 are $(m;\ 1, 1, \ldots .\ 1)$,

$$\theta_1 = \overbrace{\boxed{\begin{array}{cccccc} g_1 & 1 & 0 & 0 & 0 & \ldots \\ 0 & g_1 & 1 & 0 & 0 & \ldots \\ 0 & 0 & g_1 & 1 & 0 & \ldots \\ 0 & 0 & 0 & g_1 & 1 & \ldots \\ 0 & 0 & 0 & 0 & g_1 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{array}}}^{m} \left.\vphantom{\begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{array}}\right\} m$$

In like manner, in the subsidiary matrix θ_2 the principal diagonal is bordered on the right by unity, and the constituents in the principal diagonal are all equal to g_2, the remaining $m^2 - 2m + 1$ constituents being zero; similarly with respect to the remaining squares, θ_3, etc.

Let ϑ_1, ϑ_2, etc., denote the matrices obtained from θ by erasing successively all the constituents of θ except, first, those in the square θ_1; second, those in the square θ_2, etc. Then

$$\theta = \vartheta_1 + \vartheta_2 + \vartheta_3 + \cdots ;$$

and since the ϑ's are nilfactorial together, i. e. since $\vartheta_1 \vartheta_2 = \vartheta_2 \vartheta_1 = 0$, etc., therefore

$$\theta^2 = \vartheta_1^2 + \vartheta_2^2 + \vartheta_3^2 + \cdots\cdot,$$

$$\theta^3 = \vartheta_1^3 + \vartheta_2^3 + \vartheta_3^3 + \cdots\cdot,$$

etc., etc.

Consequently, if $\boldsymbol{f}(\theta)$ is any rational integral function of θ of which k is the term independent of θ, and if $f(\theta) = \boldsymbol{f}(\theta) - k$, then

$$\boldsymbol{f}(\theta) = f(\theta) + k = f(\vartheta_1) + f(\vartheta_2) + f(\vartheta_3) + \cdots\cdot + k.$$

Let $[r, s]$ denote the matrix whose constituents are all zero except that in the r-th row and s-th column, which is unity. We have then

$$\vartheta_1 = g_1 \Sigma_{1r}^{m} [r, r] + \Sigma_{1\ r}^{m-1} [r, r + 1],$$

$$\vartheta_2 = g_2 \Sigma_{1r}^{n} [m + r, m + r] + \Sigma_{1\ r}^{n-1} [m + r, m + r + 1],$$

$$\vartheta_3 = g_3 \Sigma_{1r}^{p} [m+n+r, m+n+r] + \Sigma_{1\ r}^{p-1} [m+n+r, m+n+r+1],$$

etc., etc.

By the laws of multiplication of the symbols $[r, s]$, viz.

$$[r, s]\,[s, t] = [r, t], \qquad [r, s]\,[t, v] = 0\ (s \neq t),$$

it follows that

$$\vartheta_1^2 = g_1^2 \Sigma_{1r}^{m} [r, r] + 2\, g_1 \Sigma_{1\ r}^{m-1} [r, r + 1] + \Sigma_{1\ r}^{m-2} [r, r + 2],$$

$$\vartheta_1^3 = g_1^3 \Sigma_{1r}^{m} [r, r] + 3\, g_1^2 \Sigma_{1\ r}^{m-1} [r, r + 1] + 3\, g_1 \Sigma_{1\ r}^{m-2} [r, r + 2] + \Sigma_{1\ r}^{m-3} [r, r + 3],$$

. . . .

$$\vartheta_1^{\mu} = g_1^{\mu} \Sigma_{1r}^{m} [r, r] + \frac{d}{d g_1} g_1^{\mu} \Sigma_{1}^{m-1} [r, r+1] + \frac{1}{2} \frac{d^2}{d g_1^2} g_1^{\mu} \Sigma_{1}^{m-2} [r, r+2] + \cdots\cdot + \frac{1}{(m-1)\,!} \frac{d^{m-1}}{d\, g_1^{m-1}} g_1^{\mu} [1, m].$$

Therefore, if $f'(g)$, $f''(g)$, etc., denote respectively the first, second, etc., differential coefficients of $f(g)$ with respect to g, then

$$f(\vartheta_1) = f(g_1)\,\Sigma_{1\,r}^{\,m}[r, r] + f'(g_1)\,\Sigma_{1\,r}^{\,m-1}[r, r+1] + \frac{f''(g_1)}{2}\,\Sigma_{1\,r}^{\,m-2}[r, r+2]$$

$$+ \frac{f'''(g_1)}{2.3}\,\Sigma_{1\,r}^{\,m-3}[r, r+3] + \cdots\cdots + \frac{f^{(m-1)}(g_1)}{(m-1)!}[1, m]$$

$$= \Sigma_{0\,s}^{\,m-1}\left\{\frac{f^{(s)}(g_1)}{s!}\,\Sigma_{1\,r}^{\,m-s}[r, r+s]\right\},$$

Similarly, we have

$$f(\vartheta_2) = f(g_2)\,\Sigma_{1\,r}^{\,n}[m+r, m+r] + f'(g_2)\,\Sigma_{1\,r}^{\,n}[m+r, m+r+1]$$

$$+ \frac{f''(g_2)}{2}\,\Sigma_{1\,r}^{\,n-2}[m+r, m+r+2] + \cdots\cdots + \frac{f^{(n-1)}(g_2)}{(n-1)!}[m+1, m+n]$$

$$= \Sigma_{0\,s}^{\,n-1}\left\{\frac{f^{(s)}(g_2)}{s!}\,\Sigma_{1\,r}^{\,n-s}[m+r, m+r+s]\right\},$$

$$f(\vartheta_3) = f(g_3)\,\Sigma_{1\,r}^{\,p}[m+n+r, m+n+r]$$

$$+ f'(g_3)\,\Sigma_{1\,r}^{\,p-1}[m+n+r, m+n+r+1]$$

$$+ \frac{f''(g_3)}{2}\,\Sigma_{1\,r}^{\,p-2}[m+n+r, m+n+r+2] + \cdots\cdots$$

$$+ \frac{f^{(p-1)}(g_3)}{(p-1)!}[m+n+1, m+n+p]$$

$$= \Sigma_{0\,s}^{\,p-1}\left\{\frac{f^{(s)}(g_3)}{s!}\,\Sigma_{1\,r}^{\,p-s}[m+n+r, m+n+r+s]\right\},$$

etc., etc.

And since $\boldsymbol{f}'(g) = f'(g)$, $\boldsymbol{f}''(g) = f''(g)$, etc., and

$$1 = \Sigma_{1\,r}^{\,m}[r, r] + \Sigma_{1\,r}^{\,n}[m+r, m+r]$$

$$+ \Sigma_{1\,r}^{\,p}[m+n+r, m+n+r] + \cdots\cdots,$$

it follows that

$$\boldsymbol{f}(\theta) = f(\theta) + k$$

$$= f(\vartheta_1) + f(\vartheta_2) + f(\vartheta_3) + \cdots\cdots + k$$

$$= \Sigma_{0\,s}^{\,m-1}\left\{\frac{\boldsymbol{f}^{(s)}(g_1)}{s!}\,\Sigma_{1\,r}^{\,m}[r, r+s]\right\}$$

$$+ \Sigma_{0\,s}^{\,n-1}\left\{\frac{\boldsymbol{f}^{(s)}(g_2)}{s!}\,\Sigma_{1\,r}^{\,n}[m+r, m+r+s]\right\}$$

$$+ \Sigma_{0\,s}^{\,p-1}\left\{\frac{\boldsymbol{f}^{(s)}(g_3)}{s!}\,\Sigma_{1\,r}^{\,p}[m+n+r, m+n+r+s]\right\} + \text{etc.}$$

Let now

$$f_1(\theta) = [(\theta - g_1)^m - (g_2 - g_1)^m]^n\,[(\theta - g_1)^m - (g_3 - g_1)^m]^p \ldots,$$

the remaining factors corresponding to the remaining latent roots; then for $g = g_1$,

$$f_1(g) = f_1(g_1) \neq 0,\ f_1'(g) = 0,\ f_1''(g) = 0,\ \ldots\, f_1^{(m-1)}(g) = 0;$$

for $g = g_2$, $f_1(g) = 0$, and likewise its first $n - 1$ differential coefficients; similarly, for $g = g_3$, $f_1(g)$ and its first $p - 1$ differential coefficients are all zero, etc. Therefore

$$\frac{f_1(\theta)}{f_1(g_1)} = \Sigma_{1r}^{m}\,[r,\ r].$$

It may be shown in like manner that

$$\frac{f_1(\theta)}{f_1(g_1)}\,(\theta - g_1) = \Sigma_{1\,r}^{m-1}\,[r,\ r + 1],$$

$$\frac{f_1(\theta)}{f_1(g_1)}\,(\theta - g_1)^2 = \Sigma_{1\,r}^{m-2}\,[r,\ r + 2],$$

. . . .

$$\frac{f_1(\theta)}{f_1(g_1)}\,(\theta - g_1)^{m-1} = [1,\ m].$$

Similarly, if

$$f_2(\theta) = [(\theta - g_2)^n - (g_1 - g_2)^n]^m\,[(\theta - g_2)^n - (g_3 - g_2)^n]^p \ldots,$$

then it may be shown in like manner that

$$\frac{f_2(\theta)}{f_2(g_2)} = \Sigma_{1r}^{n}\,[m + r,\ m + r],$$

$$\frac{f_2(\theta)}{f_2(g_2)}\,(\theta - g_2) = \Sigma_{1\,r}^{n-1}\,[m + r,\ m + r + 1],$$

$$\frac{f_2(\theta)}{f_2(g_2)}\,(\theta - g_2)^2 = \Sigma_{1\,r}^{n-2}\,[m + r,\ m + r + 2],$$

. . . .

$$\frac{f_2(\theta)}{f_2(g_2)}\,(\theta - g_2)^{n-1} = [m + 1,\ m + n],$$

etc., etc.

Now, by reference to the theorem referred to above, it will be seen that if ϕ is the most general matrix commutative with Ω, then ϕ may be written as $\omega\,\eta\,\omega^{-1}$ where

$$\eta = \left(\begin{array}{|cccc|} \boxed{\eta_1} & & & \\ & \boxed{\eta_2} & & \\ & & \boxed{\eta_3} & \\ & & & \ddots \end{array}\right),$$

in which all the constituents are zero except those in the squares η_1, η_2, etc., along the principal diagonal, and these square arrays or subsidiary matrices are of order equal respectively to the order of the subsidiary matrices θ_1, θ_2, etc., to which they severally correspond; moreover,

$$\eta_1 = \overbrace{\begin{array}{|cccccc|} \hline a_1 & b_1 & c_1 & d_1 & e_1 & \dots \\ 0 & a_1 & b_1 & c_1 & d_1 & \dots \\ 0 & 0 & a_1 & b_1 & c_1 & \dots \\ 0 & 0 & 0 & a_1 & b_1 & \dots \\ 0 & 0 & 0 & 0 & a_1 & \dots \\ \text{etc.} & \dots & \dots & \dots & \dots & \dots \\ \hline \end{array}}^{m} \left.\vphantom{\begin{array}{c} a\\a\\a\\a\\a\\a \end{array}}\right\} m$$

$$\eta_2 = \overbrace{\begin{array}{|cccccc|} \hline a_2 & b_2 & c_2 & d_2 & e_2 & \dots \\ 0 & a_2 & b_2 & c_2 & d_2 & \dots \\ 0 & 0 & a_2 & b_2 & c_2 & \dots \\ 0 & 0 & 0 & a_2 & b_2 & \dots \\ 0 & 0 & 0 & 0 & a_2 & \dots \\ \text{etc.} & \dots & \dots & \dots & \dots & \dots \\ \hline \end{array}}^{n} \left.\vphantom{\begin{array}{c} a\\a\\a\\a\\a\\a \end{array}}\right\} n$$

etc., etc.

Therefore, if ϵ_1, ϵ_2, etc., severally denote the matrices formed from η by erasing successively all its constituents except, first, those in the square η_1, second, those in the square η_2, etc., then

$$\eta = \epsilon_1 + \epsilon_2 + \cdots\cdot,$$

and

$$\epsilon_1 = a_1 \Sigma_{1}{}_{r}^{m} \big[r,\ r\big] + b_1 \Sigma_{1}{}^{m-1}_{r} \big[r,\ r+1\big] + c_1 \Sigma_{1}{}^{m-2}_{r} \big[r,\ r+2\big]$$
$$+ d_1 \Sigma_{1}{}^{m-3}_{r} \big[r,\ r+3\big] + e_1 \Sigma_{1}{}^{m-4}_{r} \big[r,\ r+4\big] + \cdots\cdot$$

$$\epsilon_2 = a_2 \Sigma_{1}{}_{r}^{n} \big[m+r,\ m+r\big] + b_2 \Sigma_{1}{}^{n-1}_{r} \big[m+r,\ m+r+1\big]$$
$$+ c_2 \Sigma_{1}{}^{n-2}_{r} \big[m+r,\ m+r+2\big] + d_2 \Sigma_{1}{}^{n-3}_{r} \big[m+r,\ m+r+3\big]$$
$$+ e_2 \Sigma_{1}{}^{n-4}_{r} \big[m+r,\ m+r+4\big] + \cdots\cdot,$$

etc., etc.

But by what has just been proved it follows that

$$\epsilon_1 = \frac{f_1(\theta)}{f_1(g_1)} \{a_1 + b_1(\theta - g_1) + c_1(\theta - g_1)^2 + d_1(\theta - g_1)^3 + e_1(\theta - g_1)^4 + \cdots\cdot\},$$

$$\epsilon_2 = \frac{f_2(\theta)}{f_2(g_2)} \{a_2 + b_2(\theta - g_2) + c_2(\theta - g_2)^2 + d_2(\theta - g_2)^3 + e_2(\theta - g_2)^4 + \cdots\cdot\},$$

etc. Consequently η is a rational integral function of θ. Let $\eta = f(\theta)$, where $f(\theta)$ denotes the most general rational integral function of θ; then, since

$$\omega \cdot f(\theta) \cdot \omega^{-1} = f(\Omega),$$

the most general matrix commutative with Ω is

$$\phi = \omega \eta \omega^{-1} = \omega \cdot f(\theta) \cdot \omega^{-1} = f(\Omega).$$

Worcester, Mass., January 2, 1892.

POSTSCRIPT.

WORCESTER, MASS., March 26, 1892.

I find that the solution of the matrical equation $\Omega \phi = \phi \Omega$ given in a note in these Proceedings, May 26, 1891, and upon which the proof of the theorem given above is based, gives the complete solution of the more general matrical equation $\Omega \phi = \phi \Omega'$ for that case in which it may be assumed that the determinant of ϕ is not null. For this case the latent roots of Ω' are identical with those of Ω, and have, respectively, the same characteristics, i. e. the nullities of successive powers of Ω' less any latent root g are respectively equal to the nullities of successive powers of $\Omega - g$. Therefore, if the canonical form of Ω is $\omega \theta \omega^{-1}$, the canonical form of Ω' is $\omega' \theta \omega'^{-1}$.* Matrices thus related are termed by Weyr matrices of the same kind.

If, then, in the equation

$$\Omega \phi = \phi \Omega',$$

Ω and Ω' are replaced by their equivalents, it becomes

$$\omega \theta \omega^{-1} \phi = \phi \omega' \theta \omega'^{-1},$$

that is,

$$\theta \, . \, \omega^{-1} \phi \omega' = \omega^{-1} \phi \omega' \, . \, \theta.$$

If, now, we put

$$\psi = \omega^{-1} \phi \omega',$$

the equation becomes

$$\theta \psi = \psi \theta.$$

But the most general solution of this equation is given in the note above referred to; the most general matrix commutative with θ is there denoted by η. Therefore, the most general solution of the matrical equation

$$\phi \Omega = \Omega' \phi$$

is

$$\phi = \omega \psi \omega'^{-1},$$

---

* For explanation of the term canonical form of a matrix, see paper referred to, these Proceedings, Vol. XXVI p. 64.

if the canonical forms of Ω and Ω' are respectively $\omega\,\theta\,\omega^{-1}$ and $\omega'\,\theta\,\omega'^{-1}$, and provided ψ is the most general matrix commutative with θ.

This solution of the equation $\Omega\,\phi = \phi\,\Omega'$, as also the above mentioned solution of the equation $\Omega\,\phi = \phi\,\Omega$, depends upon the reduction of a matrix to its canonical form. In the assumed canonical form of Ω, viz. $\omega\,\theta\,\omega^{-1}$, the matrix θ is dependent solely upon the characteristics of the latent roots of Ω. Having determined the latent roots of Ω and their characteristics, if one value of ω is known, the solution of the preceding equation $\Omega\,\phi = \phi\,\Omega'$ gives in terms of ω every solution of the problem to reduce a given matrix Ω to its canonical form. For if $\Omega' = \theta$, the most general solution of the equation

$$\Omega\,\phi = \phi\,\Omega'$$

reduces to

$$\phi = \omega\,\psi,$$

where ψ is the most general matrix commutative with θ, and ω is some particular value of ϕ satisfying the equation $\Omega = \phi\,\theta\,\phi^{-1}$. But since the solutions of this equation are all included in those of the equation $\Omega\,\phi = \phi\,\theta$, if the constants in ψ are so taken that its determinant (and consequently the determinant of ϕ) is not null, $\phi = \omega\,\psi$ is the most general solution of the problem to reduce a matrix to its canonical form.

If Ω and Ω' are matrices of the same kind,* and such that for any latent root g the increments of nullity of successive powers of $\Omega - g$ (or $\Omega' - g$) are severally equal to the nullity of $\Omega - g$ (or of $\Omega' - g$), until a power is reached whose nullity is equal to its vacuity, then the most general solution of the equation $\Omega\,\phi = \phi\,\Omega'$ can be expressed in terms of certain arbitrary matrices and of products of powers of Ω and Ω' less their latent roots. Thus, let the latent roots of Ω (and of Ω') be $g_1, g_2, \ldots. g_i$, respectively of multiplicity $m_1, m_2, \ldots. m_i$; and suppose that the nullities of $\Omega - g_1$, $\Omega - g_2$, etc. are m'_1, m'_2, etc. From the supposed equality in the increments of nullity of $\Omega - g_1$, etc., it follows that

$$\mu_1 = \frac{m_1}{m'_1}, \quad \mu_2 = \frac{m_2}{m'_2}, \ldots. \quad \mu_i = \frac{m_i}{m'_i},$$

are all integers. Let, now,

---

* That is, Ω and Ω' are such that it is possible to find a matrix χ to satisfy the equation $\Omega' = \chi\,\Omega\,\chi^{-1}$.

$$f_r(g) \equiv (\overline{g - g_r}^{\mu_r} - \overline{g_1 - g_r}^{\mu_r})^{\mu_1} (\overline{g - g_r}^{\mu_r} - \overline{g_2 - g_r}^{\mu_r})^{\mu_2} \ldots\ldots$$

$$(\overline{g - g_r}^{\mu_r} - \overline{g_{r-1} - g_r})^{\mu_r - 1} (\overline{g - g_r}^{\mu_r} - \overline{g_{r+1} - g_r}^{\mu_r})^{\mu_r + 1} \ldots\ldots$$

$$(\overline{g - g_r}^{\mu_r} - \overline{g_i - g_r}^{\mu_r})^{\mu_i},$$

$$\Omega_r = \frac{f_r(\Omega)}{f_r(g_r)}, \qquad \Omega'_r = \frac{f_r(\Omega')}{f_r(g_r)}.$$

Then the most general expression for the matrix ϕ satisfying the equation $\Omega\,\phi = \phi\,\Omega'$ is

$$\phi = \Sigma_{0\,r}^{\mu_1 - 1}\{(\Omega - g_1)^{\mu_1 - r - 1}\,\Omega_1\,\mathrm{M}_1\,\Omega'_1\,(\Omega' - g_1)^r\}$$

$$+ \Sigma_{0\,r}^{\mu_2 - 1}\{(\Omega - g_2)^{\mu_2 - r - 1}\,\Omega_2\,\mathrm{M}_2\,\Omega'_2\,(\Omega' - g_2)^r\}$$

$$+ \ldots\ldots + \Sigma_{0\,r}^{\mu_i - 1}\{(\Omega - g_i)^{\mu_i - r - 1}\,\Omega_i\,\mathrm{M}\,\Omega'_i\,(\Omega' - g_i)_r\},$$

where M_1, M_2, ... M_i, are arbitrary matrices.

If $\Omega' = \Omega$, the expression to which this expression for ϕ reduces gives the most general matrix commutative with Ω.

If $\Omega' = \theta$, and the matrices M in the above expression for ϕ are so taken (as is possible) that the determinant of ϕ is not null, then the expression for ϕ gives the most general solution of the reduction of the given matrix Ω to its canonical form.